SEQUENCE LISTING

(1) GENERAL INFORMATION:

    (i) APPLICANT: Gorski, David H.
             Walsh, Kenneth

   (ii) TITLE OF INVENTION: Growth Arrest Homeobox Gene

  (iii) NUMBER OF SEQUENCES: 4

   (iv) CORRESPONDENCE ADDRESS:
      (A) ADDRESSEE: Calfee, Halter, and Griswold
      (B) STREET: 800 Superior Avenue
      (C) CITY: Cleveland
      (D) STATE: Ohio
      (E) COUNTRY: U.S.A.
      (F) ZIP: 44114-2688

    (v) COMPUTER READABLE FORM:
      (A) MEDIUM TYPE: Floppy disk
      (B) COMPUTER: IBM PC compatible
      (C) OPERATING SYSTEM: PC-DOS/MS-DOS
      (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

   (vi) CURRENT APPLICATION DATA:
      (A) APPLICATION NUMBER:
      (B) FILING DATE:
      (C) CLASSIFICATION:

 (viii) ATTORNEY/AGENT INFORMATION:
      (A) NAME: Golrick, Mary E.
      (B) REGISTRATION NUMBER: 34829
      (C) REFERENCE/DOCKET NUMBER: 22311/00114

   (ix) TELECOMMUNICATION INFORMATION:
      (A) TELEPHONE: (216) 622-8200
      (B) TELEFAX: (216) 241-0816
      (C) TELEX: 980499

(2) INFORMATION FOR SEQ ID NO:1:

    (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 2244 base pairs
      (B) TYPE: nucleic acid
      (C) STRANDEDNESS: both
      (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: cDNA

  (iii) HYPOTHETICAL: NO

   (iv) ANTI-SENSE: NO

   (ix) FEATURE:
      (A) NAME/KEY: CDS
      (B) LOCATION: 197..1108

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GTCAAGTGTT TATACGTGCA GGAGACTGGC CGCTCGGCTC AGGACTGGGA TTAGCGGGCT
60

CTGCTCAAAC CCGCGCGGCT TTTACATTAG GAGTGAGTGG GGGAGAGTCC TAGGATTTCT
120

AGTGAAAAGT GACAGCGCTT GGTGGACTTT GGGACCTTCG TGAAGTCTTC TGCTTGGAAG
180

CTGAGACTTG CATGCC ATG GAA CAC CCC CTC TTT GGC TGC CTG CGC AGC
229
                    Met Glu His Pro Leu Phe Gly Cys Leu Arg Ser
                     1   5   10

CCC CAC GCC ACA GCG CAA GGC TTG CAC CCC TTC TCG CAG TCT TCT CTG
277
Pro His Ala Thr Ala Gln Gly Leu His Pro Phe Ser Gln Ser Ser Leu
             15   20   25

GCC CTC CAT GGA AGA TCT GAC CAC ATG TCC TAC CCC GAA CTC TCC ACA
325
Ala Leu His Gly Arg Ser Asp His Met Ser Tyr Pro Glu Leu Ser Thr
         30   35   40

TCT TCC TCG TCT TGC ATA ATC GCG GGA TAC CCC AAT GAG GAG GGC ATG
373
Ser Ser Ser Ser Cys Ile Ile Ala Gly Tyr Pro Asn Glu Glu Gly Met
         45   50   55

TTT GCC AGC CAG CAT CAC AGG GGG CAC CAC CAC CAC CAC CAC CAC CAC
421
Phe Ala Ser Gln His His Arg Gly His His His His His His His His
 60   65   70   75

CAT CAC CAC CAC CAG CAG CAG CAG CAC CAG GCT CTG CAA AGC AAC TGG
469
His His His His Gln Gln Gln Gln His Gln Ala Leu Gln Ser Asn Trp
                     80   85   90

CAC CTC CCG CAG ATG TCC TCC CCG CCA AGC GCG GCC CGG CAC AGC CTT
517
His Leu Pro Gln Met Ser Ser Pro Pro Ser Ala Ala Arg His Ser Leu
             95   100   105

TGC CTG CAG CCT GAT TCC GGA GGG CCC CCG GAG CTG GGG AGC AGC CCT
565
Cys Leu Gln Pro Asp Ser Gly Gly Pro Pro Glu Leu Gly Ser Ser Pro
             110   115   120

CCG GTC CTG TGC TCC AAC TCT TCT AGC CTG GGC TCC AGC ACC CCG ACC
613
Pro Val Leu Cys Ser Asn Ser Ser Ser Leu Gly Ser Ser Thr Pro Thr
         125   130   135

GGA GCC GCG TGC GCA CCA AGG GAT TAT GGC CGT CAA GCG CTG TCA CCC
661
Gly Ala Ala Cys Ala Pro Arg Asp Tyr Gly Arg Gln Ala Leu Ser Pro
140   145   150   155

GCA GAA GTG GAG AAG AGA AGT GGC AGC AAA AGA AAA AGC GAC AGT TCA
709
Ala Glu Val Glu Lys Arg Ser Gly Ser Lys Arg Lys Ser Asp Ser Ser
                 160   165   170

GAT TCC CAG GAA GGA AAT TAC AAG TCA GAA GTG AAC AGC AAA CCT AGG
757
Asp Ser Gln Glu Gly Asn Tyr Lys Ser Glu Val Asn Ser Lys Pro Arg
                 175   180   185

AGG GAA AGA ACA GCT TTC ACC AAA GAG CAA ATC AGA GAA CTT GAG GCA

SEQUENCE LISTING

(1) GENERAL INFORMATION:

    (i) APPLICANT: Gorski, David H.
               Walsh, Kenneth

    (ii) TITLE OF INVENTION: Growth Arrest Homeobox Gene

   (iii) NUMBER OF SEQUENCES: 4

    (iv) CORRESPONDENCE ADDRESS:
        (A) ADDRESSEE: Calfee, Halter, and Griswold
        (B) STREET: 800 Superior Avenue
        (C) CITY: Cleveland
        (D) STATE: Ohio
        (E) COUNTRY: U.S.A.
        (F) ZIP: 44114-2688

    (v) COMPUTER READABLE FORM:
        (A) MEDIUM TYPE: Floppy disk
        (B) COMPUTER: IBM PC compatible
        (C) OPERATING SYSTEM: PC-DOS/MS-DOS
        (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

    (vi) CURRENT APPLICATION DATA:
        (A) APPLICATION NUMBER:
        (B) FILING DATE:
        (C) CLASSIFICATION:

   (viii) ATTORNEY/AGENT INFORMATION:
        (A) NAME: Golrick, Mary E.
        (B) REGISTRATION NUMBER: 34829
        (C) REFERENCE/DOCKET NUMBER: 22311/00114

    (ix) TELECOMMUNICATION INFORMATION:
        (A) TELEPHONE: (216) 622-8200
        (B) TELEFAX: (216) 241-0816
        (C) TELEX: 980499

(2) INFORMATION FOR SEQ ID NO:1:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 2244 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: both
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

   (iii) HYPOTHETICAL: NO

    (iv) ANTI-SENSE: NO

    (ix) FEATURE:
        (A) NAME/KEY: CDS
        (B) LOCATION: 197..1108

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GTCAAGTGTT TATACGTGCA GGAGACTGGC CGCTCGGCTC AGGACTGGGA TTAGCGGGCT
60

CTGCTCAAAC CCGCGCGGCT TTTACATTAG GAGTGAGTGG GGGAGAGTCC TAGGATTTCT
120

AGTGAAAAGT GACAGCGCTT GGTGGACTTT GGGACCTTCG TGAAGTCTTC TGCTTGGAAG
180

CTGAGACTTG CATGCC ATG GAA CAC CCC CTC TTT GGC TGC CTG CGC AGC
229
              Met Glu His Pro Leu Phe Gly Cys Leu Arg Ser
               1   5   10

CCC CAC GCC ACA GCG CAA GGC TTG CAC CCC TTC TCG CAG TCT TCT CTG
277
Pro His Ala Thr Ala Gln Gly Leu His Pro Phe Ser Gln Ser Ser Leu
         15   20   25

GCC CTC CAT GGA AGA TCT GAC CAC ATG TCC TAC CCC GAA CTC TCC ACA
325
Ala Leu His Gly Arg Ser Asp His Met Ser Tyr Pro Glu Leu Ser Thr
         30   35   40

TCT TCC TCG TCT TGC ATA ATC GCG GGA TAC CCC AAT GAG GAG GGC ATG
373
Ser Ser Ser Ser Cys Ile Ile Ala Gly Tyr Pro Asn Glu Glu Gly Met
         45   50   55

TTT GCC AGC CAG CAT CAC AGG GGG CAC CAC CAC CAC CAC CAC CAC CAC
421
Phe Ala Ser Gln His His Arg Gly His His His His His His His His
 60   65   70   75

CAT CAC CAC CAC CAG CAG CAG CAG CAC CAG GCT CTG CAA AGC AAC TGG
469
His His His His Gln Gln Gln Gln His Gln Ala Leu Gln Ser Asn Trp
                 80   85   90

CAC CTC CCG CAG ATG TCC TCC CCG CCA AGC GCG GCC CGG CAC AGC CTT
517
His Leu Pro Gln Met Ser Ser Pro Pro Ser Ala Ala Arg His Ser Leu
             95   100   105

TGC CTG CAG CCT GAT TCC GGA GGG CCC CCG GAG CTG GGG AGC AGC CCT
565
Cys Leu Gln Pro Asp Ser Gly Gly Pro Pro Glu Leu Gly Ser Ser Pro
         110   115   120

CCG GTC CTG TGC TCC AAC TCT TCT AGC CTG GGC TCC AGC ACC CCG ACC
613
Pro Val Leu Cys Ser Asn Ser Ser Ser Leu Gly Ser Ser Thr Pro Thr
         125   130   135

GGA GCC GCG TGC GCA CCA AGG GAT TAT GGC CGT CAA GCG CTG TCA CCC
661
Gly Ala Ala Cys Ala Pro Arg Asp Tyr Gly Arg Gln Ala Leu Ser Pro
140   145   150   155

GCA GAA GTG GAG AAG AGA AGT GGC AGC AAA AGA AAA AGC GAC AGT TCA
709
Ala Glu Val Glu Lys Arg Ser Gly Ser Lys Arg Lys Ser Asp Ser Ser
                 160   165   170

GAT TCC CAG GAA GGA AAT TAC AAG TCA GAA GTG AAC AGC AAA CCT AGG
757
Asp Ser Gln Glu Gly Asn Tyr Lys Ser Glu Val Asn Ser Lys Pro Arg
                 175   180   185

AGG GAA AGA ACA GCT TTC ACC AAA GAG CAA ATC AGA GAA CTT GAG GCA

805
Arg Glu Arg Thr Ala Phe Thr Lys Glu Gln Ile Arg Glu Leu Glu Ala
        190   195   200

GAG TTC GCC CAT CAT AAC TAT CTG ACC AGA CTG AGA AGA TAT GAG ATA
853
Glu Phe Ala His His Asn Tyr Leu Thr Arg Leu Arg Arg Tyr Glu Ile
        205   210   215

GCG GTG AAC CTA GAC CTC ACT GAA AGA CAG GTG AAA GTG TGG TTC CAG
901
Ala Val Asn Leu Asp Leu Thr Glu Arg Gln Val Lys Val Trp Phe Gln
220   225   230   235

AAC AGG AGA ATG AAG TGG AAG CGG GTC AAG GGG GGA CAA CAA GGA GCT
949
Asn Arg Arg Met Lys Trp Lys Arg Val Lys Gly Gly Gln Gln Gly Ala
            240   245   250

GCA GCC CGA GAA AAG GAA CTG GTG AAT GTG AAA AAG GGA ACA CTT CTT
997
Ala Ala Arg Glu Lys Glu Leu Val Asn Val Lys Lys Gly Thr Leu Leu
            255   260   265

CCA TCA GAG CTG TCA GGA ATT GGT GCA GCC ACC CTC CAG CAG ACA GGG
1045
Pro Ser Glu Leu Ser Gly Ile Gly Ala Ala Thr Leu Gln Gln Thr Gly
        270   275   280

GAC TCA CTA GCA AAT GAC GAC AGT CGC GAT AGT GAC CAC AGC TCT GAG
1093
Asp Ser Leu Ala Asn Asp Asp Ser Arg Asp Ser Asp His Ser Ser Glu
        285   290   295

CAC GCA CAC TTA TGATACATAC AGAGACCAGC TCCGTTCTCA GGAAAGCACC
1145
His Ala His Leu
300

ATTGTGATGG CAAATCTCAC CCAAACATCG TTTACATGGC AGATGACTGT GGCAGTGTTG
1205

CTTAATATAA TTAAACGCAG GCATCTCAAG TCTGTTTCTC ATGATTGATA GAAGGTTTAC
1265

ACTAAGTGCC TCTTATTGAA GATGCTTCCA CAGTGAAATT GGAGAAAGTG AACATATCTA
1325

AATATACTTG TTCCTTATAT GACAGAGAGG GAGATGAATG TTTGCTTTGG CTTGCACTGA
1385

AAATTAAATT GCTACCAAGA GCAAACTCGG TAAGACATTT TGACTCAAGT TGTCTCCAGA
1445

GTGAAGATGT TATAGAAATG CTTTGAACAT TCCAGTTGTA CCAGGTCATG TGTGTGACAC
1505

TGGGCAGGTA TTTGCTTTTG CTTGCACTGA AACTTAAACT GCTATCAAGT TAACCCATGA
1565

AATAGTTTAT CTTGAACAGC CACAGTGCCT GAAATCACCA AGTGGATATA AAATGAACTG
1625

AAATTCTGTA TATATTACTC CTAAGTCATT TTCCTGTCTT CACTAATTTT AGCAAATGCA
1685

TTCATATTAG CTGATGAAAA TAGGCTTTCC CGTGGACAAA TGCAGCCAGC TTCTTGTATT

1745

TTTATACATT TTTTTGTCAG TCAGAGACAT CAGTATGTGC TTACTTGTGT TCAAGTAGAG
1805

GAAATGCAGT AGAGTCTGAT AGGACATATT CTTGGTACCA CAGACAAAAC AAATCTTCTG
1865

TTGCATTGAC TATCAACTGC TGCAGATACA TTAGAGAACA CACCTAGCCC CCCTCCAGCC
1925

TCCCTCTGTT ATCGCTCGAA GACATTAGCG TCATAGGCAA GTAGTTACCT TGCCAAATGA
1985

GTCTTGTGTG GCAGATGTCT GATTTTGTAT CTTTAAACTG TTAATGGTAT GTGTCTGCTT
2045

CAGTTAACAG GGAAAAAGAT TTCTTCCTCA TTGTTTATGA TACAAAACCC AAGTGCCAAA
2105

CAAAGCTAGT TCTTCAAGGG ATAGATGAGA AACTGAATGT CTGACAAGTA GACTCAGCGA
2165

AAATACATTA TTTTTCAGAG GCTGTGTATT CATGCAGTAC AAGTCCTTGT ATTTTGTAAA
2225

AAAAAAAGTT AAATAAATG
2244

(2) INFORMATION FOR SEQ ID NO:2:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 303 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: protein

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Glu His Pro Leu Phe Gly Cys Leu Arg Ser Pro His Ala Thr Ala
 1   5   .10   15

Gln Gly Leu His Pro Phe Ser Gln Ser Ser Leu Ala Leu His Gly Arg
             20   25   30

Ser Asp His Met Ser Tyr Pro Glu Leu Ser Thr Ser Ser Ser Ser Cys
         35   40   45

Ile Ile Ala Gly Tyr Pro Asn. Glu Glu Gly Met Phe Ala Ser Gln His
     50   55   60

His Arg Gly His His His His His His His His His His His Gln
     65   70   75   80

Gln Gln Gln His Gln Ala Leu Gln Ser Asn Trp His Leu Pro Gln Met
             85   90   95

Ser Ser Pro Pro Ser Ala Ala Arg His Ser Leu Cys Leu Gln Pro Asp
             100   105   110

Ser Gly Gly Pro Pro Glu Leu Gly Ser Ser Pro Pro Val Leu Cys Ser
         115   120   125

Asn Ser Ser Ser Leu Gly Ser Ser Thr Pro Thr Gly Ala Ala Cys Ala
     130   135   140

Pro Arg Asp Tyr Gly Arg Gln Ala Leu Ser Pro Ala Glu Val Glu Lys
145   150   155   160

Arg Ser Gly Ser Lys Arg Lys Ser Asp Ser Ser Asp Ser Gln Glu Gly
                165   170   175

Asn Tyr Lys Ser Glu Val Asn Ser Lys Pro Arg Arg Glu Arg Thr Ala
                180   185   190

Phe Thr Lys Glu Gln Ile Arg Glu Leu Glu Ala Glu Phe Ala His His
                195   200   205

Asn Tyr Leu Thr Arg Leu Arg Arg Tyr Glu Ile Ala Val Asn Leu Asp
            210   215   220

Leu Thr Glu Arg Gln Val Lys Val Trp Phe Gln Asn Arg Arg Met Lys
225   230   235   240

Trp Lys Arg Val Lys Gly Gly Gln Gln Gly Ala Ala Ala Arg Glu Lys
                245   250   255

Glu Leu Val Asn Val Lys Lys Gly Thr Leu Leu Pro Ser Glu Leu Ser
            260   265   270

Gly Ile Gly Ala Ala Thr Leu Gln Gln Thr Gly Asp Ser Leu Ala Asn
        275   280   285

Asp Asp Ser Arg Asp Ser Asp His Ser Ser Glu His Ala His Leu
        290   295   300

(2) INFORMATION FOR SEQ ID NO:3:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 941 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: both
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (iii) HYPOTHETICAL: NO

    (iv) ANTI-SENSE: NO

    (ix) FEATURE:
        (A) NAME/KEY: CDS
        (B) LOCATION: 33..941

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GTCTTCTACC TGGAACCCGA AACTTGCATG CT ATG GAA CAC CCG CTC TTT GGC
53
                                   Met Glu His Pro Leu Phe Gly
                                   1   5

TGC CTG CGC AGC CCT CAC GCC ACG GCG CAA GGC TTG CAC CCG TTC TCC
101
Cys Leu Arg Ser Pro His Ala Thr Ala Gln Gly Leu His Pro Phe Ser
            10   15   20

CAA TCC TCT CTC GCC CTC CAT GGA AGA TCT GAC CAT ATG TCT TAC CCC
149
Gln Ser Ser Leu Ala Leu His Gly Arg Ser Asp His Met Ser Tyr Pro
        25   30   35

```
GAG CTC TCT ACT TCT TCC TCA TCT TGC ATA ATC GCG GGA TAC CCC AAC
197
Glu Leu Ser Thr Ser Ser Ser Ser Cys Ile Ile Ala Gly Tyr Pro Asn
 40              45              50                      55

GAA GAG GAC ATG TTT GCC AGC CAG CAT CAC AGG GGG CAC CAC CAC CAC
245
Glu Glu Asp Met Phe Ala Ser Gln His His Arg Gly His His His His
            60              65              70

CAC CAC CAC CAT CAC CAC CAT CAG CAG CAG CAG CAC CAG GCT CTG CAA
293
His His His His His His His Gln Gln Gln Gln His Gln Ala Leu Gln
        75              80              85

ACC AAC TGG CAC CTC CCG CAG ATG TCT TCC CCA CCG AGT GCG GCT CGG
341
Thr Asn Trp His Leu Pro Gln Met Ser Ser Pro Pro Ser Ala Ala Arg
        90              95              100

CAT AGC CTC TGC CTC CAG CCC GAC TCT GGA GGG CCC CCA GAG TTG GGG
389
His Ser Leu Cys Leu Gln Pro Asp Ser Gly Gly Pro Pro Glu Leu Gly
        105             110             115

AGC AGC CCG CCC GTC CTG TGC TCC AAC TCT TCC AGC TTG GGC TCC AGC
437
Ser Ser Pro Pro Val Leu Cys Ser Asn Ser Ser Ser Leu Gly Ser Ser
120              125             130                     135

ACC CCG ACT GGG GCC GCG TGC GCG CCG GGG GAC TAC GGC CGC CAG GCA
485
Thr Pro Thr Gly Ala Ala Cys Ala Pro Gly Asp Tyr Gly Arg Gln Ala
            140                 145             150

CTG TCA CCT GCG GAG GCG GAG AAG CGA AGC GGC GGC AAG AGG AAA AGC
533
Leu Ser Pro Ala Glu Ala Glu Lys Arg Ser Gly Gly Lys Arg Lys Ser
            155             160             165

GAC AGC TCA GAC TCC CAG GAA GGA AAT TAC AAG TCA GAA GTC AAC AGC
581
Asp Ser Ser Asp Ser Gln Glu Gly Asn Tyr Lys Ser Glu Val Asn Ser
            170             175             180

AAA CCC AGG AAA GAA AGG ACA GCA TTT ACC AAA GAG CAA ATC AGA GAA
629
Lys Pro Arg Lys Glu Arg Thr Ala Phe Thr Lys Glu Gln Ile Arg Glu
            185             190             195

CTT GAA GCA GAA TTT GCC CAT CAT AAT TAT CTC ACC AGA CTG AGG CGA
677
Leu Glu Ala Glu Phe Ala His His Asn Tyr Leu Thr Arg Leu Arg Arg
200             205             210             215

TAC GAG ATA GCA GTG AAT CTG GAT CTC ACT GAA AGA CAG GTA AAA GTC
725
Tyr Glu Ile Ala Val Asn Leu Asp Leu Thr Glu Arg Gln Val Lys Val
                220             225             230

TGG TTC CAA AAC AGG CGG ATG AAG TGG AAG AGG GTA AAG GGT GGA CAG
773
Trp Phe Gln Asn Arg Arg Met Lys Trp Lys Arg Val Lys Gly Gly Gln
            235             240             245

CAA GGA GCT GCG GCT CGG GAA AAG GAA CTG GTG AAT GTG AAA AAG GGA
821
```

```
Gln Gly Ala Ala Ala Arg Glu Lys Glu Leu Val Asn Val Lys Lys Gly
        250             255             260

ACA CTT CTC CCA TCA GAG CTG TCG GGA ATT GGT GCA GCC ACC CTC CAG
869
Thr Leu Leu Pro Ser Glu Leu Ser Gly Ile Gly Ala Ala Thr Leu Gln
        265             270             275

CAA ACA GGG GAC TCT ATA GCA AAT GAA GAC AGT CAC GAC AGT GAC CAC
917
Gln Thr Gly Asp Ser Ile Ala Asn Glu Asp Ser His Asp Ser Asp His
280             285             290             295

AGC TCA GAG CAC GCC CAC CTC TGA
941
Ser Ser Glu His Ala His Leu
                300
```

(2) INFORMATION FOR SEQ ID NO:4:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 302 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Met Glu His Pro Leu Phe Gly Cys Leu Arg Ser Pro His Ala Thr Ala
1               5               10              15

Gln Gly Leu His Pro Phe Ser Gln Ser Ser Leu Ala Leu His Gly Arg
            20              25              30

Ser Asp His Met Ser Tyr Pro Glu Leu Ser Thr Ser Ser Ser Ser Cys
        35              40              45

Ile Ile Ala Gly Tyr Pro Asn Glu Glu Asp Met Phe Ala Ser Gln His
        50              55              60

His Arg Gly His His His His His His His His His His Gln Gln
65              70              75              80

Gln Gln His Gln Ala Leu Gln Thr Asn Trp His Leu Pro Gln Met Ser
            85              90              95

Ser Pro Pro Ser Ala Ala Arg His Ser Leu Cys Leu Gln Pro Asp Ser
            100             105             110

Gly Gly Pro Pro Glu Leu Gly Ser Ser Pro Pro Val Leu Cys Ser Asn
            115             120             125

Ser Ser Ser Leu Gly Ser Ser Thr Pro Thr Gly Ala Ala Cys Ala Pro
        130             135             140

Gly Asp Tyr Gly Arg Gln Ala Leu Ser Pro Ala Glu Ala Glu Lys Arg
145             150             155             160

Ser Gly Gly Lys Arg Lys Ser Asp Ser Ser Asp Ser Gln Glu Gly Asn
            165             170             175

Tyr Lys Ser Glu Val Asn Ser Lys Pro Arg Lys Glu Arg Thr Ala Phe
            180             185             190

Thr Lys Glu Gln Ile Arg Glu Leu Glu Ala Glu Phe Ala His His Asn
```

```
                 195                    200                    205
    Tyr Leu Thr Arg Leu Arg Arg Tyr Glu Ile Ala Val Asn Leu Asp Leu
        210                    215                    220
 5
    Thr Glu Arg Gln Val Lys Val Trp Phe Gln Asn Arg Arg Met Lys Trp
    225             230                    235                    240

10  Lys Arg Val Lys Gly Gly Gln Gln Gly Ala Ala Ala Arg Glu Lys Glu
                    245                    250                    255

    Leu Val Asn Val Lys Lys Gly Thr Leu Leu Pro Ser Glu Leu Ser Gly
                    260                    265                    270
    y
15  Ile Gly Ala Ala Thr Leu Gln Gln Thr Gly Asp Ser Ile Ala Asn Glu
            275                    280                    285

    Asp Ser His Asp Ser Asp His Ser Ser Glu His Ala His Leu
        290                    295                    300
```